Aus der Poliklinik und Klinik für Zahn-, Mund- und Kieferkrankheiten
(Direktor: Prof. Dr. H. Böttger)
Westdeutsche Kieferklinik der Universität Düsseldorf
Abteilung für Prothetik und Defektprothetik

Zur Qualitätsprüfung von Dentalprodukten insbesondere auf Druckfestigkeit

Dissertation
zur Erlangung des Grades eines Doktors der Zahnmedizin
der Medizinischen Fakultät der Universität Düsseldorf

vorgelegt von
Peter Hessel
1981

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41536504_0095

Aus der Poliklinik und Klinik für Zahn-, Mund- und Kieferkrankheiten
(Direktor: Prof. Dr. H. Böttger)

Westdeutsche Kieferklinik der Universität Düsseldorf
Abteilung für Prothetik und Defektprothetik

Zur Qualitätsprüfung von Dentalprodukten insbesondere auf Druckfestigkeit

Dissertation

zur Erlangung des Grades eines Doktors der Zahnmedizin
der Medizinischen Fakultät der Universität Düsseldorf

vorgelegt von

Peter Hessel

1981

Gedruckt mit Genehmigung der Medizinischen Fakultät
der Universität Düsseldorf

Der Dekan der Medizinischen Fakultät der
Universität Düsseldorf
gez. Prof. Dr. Brüster

Der Referent
gez. Prof. Dr. Rehberg

Die Korreferenten
gez. Prof. Dr. Klinger
gez. Priv.-Doz. Dr. Stüttgen
gez. Prof. Dr. med., Dipl. Math. Repges

Inhaltsverzeichnis

	Seite
1. Einleitung	1
2. Grundzüge der Statistik	2
2.1. Fehler	2
2.2. Statistische Symbole und ihre Bedeutung	4
2.3. Grundregeln statistischer Arbeitsweise	7
2.3.1. Ergebnisangabe	7
2.3.2. Prüfung auf Gaußverteilung	8
2.3.3. Ausreißer-Tests	10
2.3.4. Zur Eleminierung von Ausreißern durch Prozentgrenzen	11
2.3.5. Ringversuche	14
3. Allgemeine Probleme der zahnärztlichen Materialprüfung	17
4. Normvorschriften zur Prüfung von Dental-Werkstoffen	20
4.1. Auswertungsvorschriften bei Druck- und Biege-Festigkeitsprüfungen in Normblättern	22
4.2. Interpretation der Auswertungsvorschriften	26

4.3. Negativbeispiel einer möglichen Meßreihe und Diskussion 28
4.4. Problematik des positiven Fehlers 31

5. Vorschläge 33
5.1. Ergebnisangabe in Standardkenndaten ? 33
5.2. Voraussetzungen an einen praktikablen Lösungsvorschlag 35
5.3. Lösungsvorschlag 37
5.4. Konsequenzen 40

6. Schluß 45

Anhang 46

Literaturverzeichnis 49

1. Einleitung

Parallel zum Vordringen der elektronischen Datenverarbeitung gewann ein Teilgebiet der Mathematik immer mehr an Bedeutung : die Statistik. Obwohl die statistischen Verfahren sehr viel älter als die EDV sind, erlaubt es jetzt der Einsatz von Rechnern, relativ komplizierte Tests ohne großen Zeitaufwand durchzuführen.

Das Ziel der Statistik ist die richtige, unverfälschende Informationskonzentration. Unverfälschend deshalb, weil mit Hilfe der Statistik auch "Ergebnisse" erzeugt werden können, die aus den Einzeldaten bei objektiver Betrachtungsweise nicht abgeleitet werden dürften.

Da in unserem Fachgebiet, der Zahnmedizin, auch geforscht und geprüft wird, sind statistische Methoden auch hier ohne weiteres anwendbar. In neueren wissenschaftlichen Arbeiten erfolgt dies mehr und mehr. Aber in den nationalen und internationalen Normen für Dentalmaterialien wird bisher praktisch nicht danach verfahren, wie noch gezeigt wird. Auf welche Art und Weise jedoch zum Beispiel die Einzelergebnisse bei der Materialprüfung ausgewertet werden könnten, wird im folgenden Bestandteil meiner Untersuchungen sein.

2. Grundzüge der Statistik

2.1. Fehler

Es sollte vorausgeschickt werden, daß es generell als unmöglich angesehen werden muß, den wahren Wert einer Größe (Länge, Temperatur, Druckfestigkeit u.s.w.) zu bestimmen. Bei jeder Messung dieser Größe treten Fehler auf. Man unterscheidet vereinfachend systematische und zufällige Fehler.

Systematische Fehler treten innerhalb einer Meßreihe bei jeder Prüfung mit gleichbleibendem Betrag und Vorzeichen auf. Dies können sein: schlecht oder ungeeichte Meßinstrumente, zu hohe oder zu niedrige Raumtemperatur, ungenaue Stoppuhren, veränderte Versuchsanordnungen, konstante Verarbeitungsfehler beim Herstellen der Prüfkörper etc. All diese Fehler führen in jedem Einzelergebnis zu einer konstanten Abweichung. Der systematische Fehler kann bekannt sein und sollte dann im Ergebnis berücksichtigt werden.

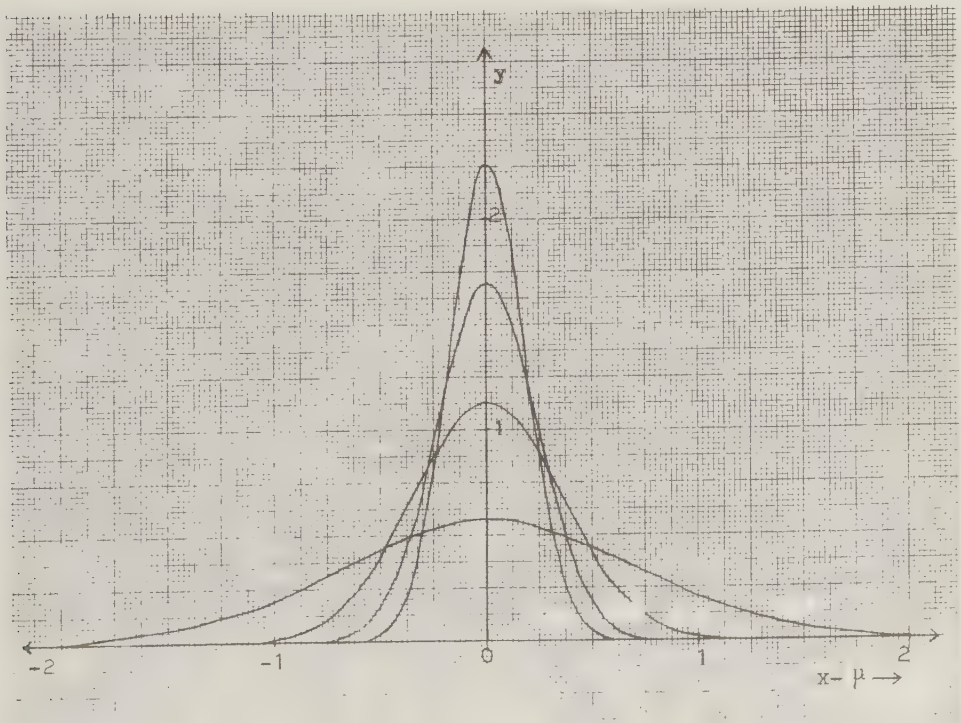
Dagegen treten zufällige Fehler nicht mit konstantem Wert und Vorzeichen auf. Sie sind nicht wie systematische Fehler vollständig vermeidbar. Durch erhöhte Sorgfalt bei den Messungen lassen sich ihre Auswirkungen lediglich mildern. Beispiele sind: Ableseungenauigkeiten auf einer großen Skala, minimale Temperaturschwankungen, zufällig auftretende Reibungskräfte in der Versuchsanordnung oder in den Meßgeräten, unterschiedlich hohe Porösitäten der Prüfkörper auf Grund schwierig konstant einhaltbarer Materialverarbeitungsvorschriften usw.

Treten nur rein zufällige Fehler auf, verteilen sich die Einzelwerte um den Populationsmittelwert μ gemäß der Gauß'schen Fehlerverteilungsfunktion :

$$y(x) = \frac{1}{\sqrt{2}\,\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Die Symbole werden im nächsten Kapitel erläutert.

Grafik 1:



2.2. Statistische Symbole und ihre Bedeutung

$x_1, x_2, x_3, \ldots x_i, \ldots x_n$
stellen die Einzelwerte einer Meßreihe dar.

n gibt die Anzahl der Einzelmessungen an.

$\bar{x}$ (gesprochen: x quer) ist der Stichprobenmittelwert. Er ist das arithmetische Mittel der Einzelwerte nach der Formel:

$$\bar{x} = \frac{x_1 + x_2 + \ldots\ldots x_n}{n} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$

μ (gesprochen: mü) ist der Populationsmittelwert. Unter einer Population würden wir im materialkundlichen Sinne die Gesamtheit der Produktionseinheiten verstehen. μ würde berechenbar, wenn man keine Stichproben aus der Produktion entnehmen würde, sondern jede produzierte Einheit (also z.B. jede produzierte Portion Amalgam) prüfen und deren Ergebnisse mitteln würde. Da dies nicht vertretbar ist, werden aus der "Population" Stichproben entnommen. Nach der Prüfung der Stichproben läßt sich durch deren Stichprobenmittelwert $\bar{x}$ der Populationsmittelwert μ abschätzen. Je größer n, die Anzahl der Einzelmessungen, ist, desto genauer ist diese Schätzung von μ.

P Die statistische Sicherheit P gibt an, mit welcher Wahrscheinlichkeit falsche Aussagen vermieden werden. P kann willkürlich festgesetzt werden. Im allgemeinen werden lediglich 3 Werte für verschiedene Anwendungsbereiche benutzt :

P= 95% bei einfachen Produktbeurteilungen
P= 99% bei wissenschaftlichen Arbeiten
P= 99,9% in besonders folgeträchtigen Situationen

Die Wahrscheinlichkeit, mit der falsche Aussagen gemacht werden, trägt das Symbol p. Beträgt z.B. P= 95%, so gilt p= 5%.

f , der Freiheitsgrad, ist eine Zahl, die je nach Test um 1 , 2 oder mehr kleiner ist als n.

$t_{f,P}$, der Student-Faktor, ist ein statistischer Tabellenwert, abhängig von f und P.
(s. Tabelle I im Anhang)

s^2 , die Stichprobenvarianz, ist ein Schätzwert für die Varianz σ^2.

$$s^2 = \frac{\sum_{i=1}^{n} (\bar{x} - x_i)^2}{n - 1}$$

Die Schätzung ist um so genauer, je höher die Anzahl der Einzelmessungen ist. Die positive Wurzel aus der Stichprobenvarianz s^2 wird als Stichprobenstandardabweichung s bezeichnet.

σ^2 ist die Populationsvarianz. In den für uns relevanten Fällen der Materialprüfung ist σ^2 jedoch unbekannt und muß mittels der Stichprobenvarianz s^2 geschätzt werden.

T ist das Symbol für den Streubereich der Einzelwerte. Er leitet sich ab nach der Formel :

$$T = s \cdot t_{n-1,P}$$

$T_{\bar{x}}$ gibt einen Bereich an von $\bar{x}-T_{\bar{x}}$ bis $\bar{x}+T_{\bar{x}}$, der den wirklichen Mittelwert μ mit der Wahrscheinlichkeit P überdeckt. Jener Bereich wird als Konfidenzintervall (Vertrauensbereich) bezeichnet.

$$T_{\bar{x}} = \frac{s \cdot t_{n-1,P}}{\sqrt{n}}$$

Es ist auch möglich, den Vertrauensbereich lediglich einseitig zu begrenzen. Unter Zuhilfenahme der t-Werte aus Tabelle II läßt sich mit oben genannter Formel ein $T_{\bar{x}}$ berechnen, welches zu folgender Aussage berechtigt : Mit der Wahrscheinlichkeit P liegt $\bar{x}-T_{\bar{x}}$ unterhalb des Populationsmittelwertes μ.

2.3. Grundregeln statistischer Arbeitsweise

Neben der Verwendung der vorgestellten Bezeichnungen, Symbole und Formeln gibt es Grundregeln, denen bei statistischen Untersuchungen Beachtung geschenkt werden muß.

2.3.1. Ergebnisangabe

Die optimale Ergebnisangabe würde es sein, alle Einzelwerte der Prüfung anzugeben. Liegen jedoch relativ viele Einzelwerte vor, würde die Ergebnisangabe sehr kompliziert werden. Eine Zusammenfassung der Information wird notwendig!

Die Informationskonzentration ist meistens sehr schnell durch die Bildung des arithmetischen Mittelwertes $\bar{x}$ erreicht. Wie exakt das dafür verwendete Prüfungsverfahren war, wie stark die Streuung der Einzelwerte gewesen ist, wieviel Einzelmessungen überhaupt durchgeführt worden sind, das alles erfahren wir erst, wenn das Ergebnis in den sogenannten Standardkenndaten angegeben wird (Kaiser/Gottschalk) :

Ergebnis: $\bar{x} \pm T$ (Einheit) ; ($\pm s$; P% ; n)

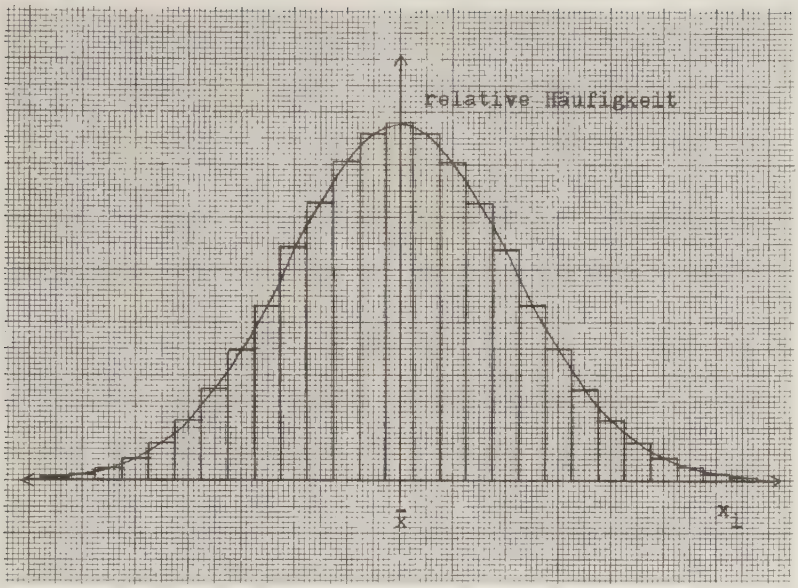
Mit ihnen erst können wir die Angaben auf ihre Richtigkeit bzw. Aussagekraft beurteilen. Jeder mit der Statistik vertraute Leser erfährt aus diesen Daten nun, wie die Prüfung verlaufen ist, ohne sich mit einer vielleicht relativ großen Anzahl von Einzelmessungen

beschäftigen zu müssen. Wenn jedoch, wie so oft, lediglich $\bar{x}$ als Ergebnis angegeben wird, ist eine objektive Beurteilung nicht mehr möglich. Z.B. kann ein Mittelwert $\bar{x} = 10$ aus den Einzelwerten 5 und 15, aber auch aus 9,99 und 10,01 gebildet werden. Beim ersten Beispiel ist der Mittelwert zwar genauso richtig berechnet wie beim zweiten Beispiel, eine Aussagekraft besitzt der zuerst ermittelte Wert aber kaum.

2.3.2. Prüfung auf Gaußverteilung

Um ein vorhandenes Datenmaterial zur Grundlage statistischer Untersuchungen machen zu können, muß es sich nach der Gaußschen Funktion verteilen. Dies liegt vor, wenn alle Abweichungen vom Mittelwert zufälliger Natur sind. Zur Überprüfung der Annahme, es handle sich um eine Gaußverteilung, nimmt man eine sogenannte-Klassierung vor, d.h.,die Einzelwerte werden je nach Größe in bestimmte Gruppen (Klassen) eingeteilt. Nun wird die Anzahl der Einzelwerte je Klasse bestimmt.
Grafisch muß sich nun ein Anhalt für eine Gauß-Verteilung ergeben. Eine deutliche Glockenform zu erkennen ist aber erst bei n, die jenseits von 30 Einzelmessungen liegen, möglich. Deshalb wird meistens auf die Prüfung des Rohmaterials verzichtet und keine Klassierung durchgeführt. Vielmehr werden mit Hilfe sogenannter Ausreißer-Tests überzufällig stark mit Fehlern behaftete Einzelmessungen entdeckt und danach aus der Meßreihe enfernt.

Grafik 2:



Jeder Block entspricht einer Klasse; die Höhe des Blocks gibt die Anzahl der Einzelwerte x_i pro Klasse an.

2.3.3. Ausreißer-Tests

Ausreißer entstehen z.B. durch grobe Unaufmerksamkeiten der Prüfer und können somit nicht mehr als rein zufällig angesehen werden. Ihre Eleminierung ist demzufolge legitim.

Der bei Kaiser/Gottschalk erwähnte Nalimov-Test kommt aus hier nicht weiter zu erläuternden Gründen nicht in Frage. In der DIN 53 804 Teil 1 werden die Tests nach Dixon und Grubbs empfohlen. Für Stichprobenumfänge $n \leq 29$ ist der Dixon-Test relevant.

Beim Dixon-Test werden die Einzelwerte nach aufsteigender Größe geordnet : $x_{(1)} \leq x_{(2)} \leq x_{(3)} \cdots\cdots x_{(n)}$ Nun kann mit Hilfe der in Tabelle III angegebenen Formeln und Schwellenwerten geprüft werden, ob $x_{(1)}$ oder $x_{(n)}$ als Ausreißer angesehen werden kann. Nach der in Abhängigkeit von n zu wählenden Formel wird ein Prüfwert errechnet. Übersteigt dieser Prüfwert den entsprechenden Schwellenwert, so darf dieser Einzelwert als Ausreißer betrachtet und eleminiert werden. Die erneute Anwendung des Tests auf die verbleibenden Einzelwerte ist nicht zulässig.

2.3.4. Zur Eleminierung von Ausreißern durch Prozentgrenzen

Gelegentlich wird in den Auswertungsvorschriften der Materialprüfung die Anweisung gegeben, Einzelwerte zu eleminieren, die um einen bestimmten Prozentsatz vom Stichprobenmittelwert abweichen. Konkrete Fälle dazu werden imKapitel 4.1. gezeigt.

An Hand von 2 Beispielen soll demonstriert werden, daß diese Methode nicht zulässig ist. Die in den Beispielen zum Ausschluß des Einzelwertes führende Prozentgrenze sei mit 15% festgesetzt.

Beispiel I :

$x_{(1)} = 84$; $x_{(2)} = 102$; $x_{(3)} = 104$; $x_{(4)} = 104$; $x_{(5)} = 106$

$\bar{x} = 100$

$s = 9{,}06$

$T_{\bar{x}} = 11{,}26$ (bei beidseitiger Abgrenzung und P=95%)

$x_{(1)}$ wäre nach der 15%-Klausel zu eleminieren. Das neue, korrigierte Ergebnis (n=4) würde lauten :

$\bar{x} = 104$

$s = 1{,}63$

$T_{\bar{x}} = 2{,}596$ (bei beidseitiger Abgrenzung und P=95%)

nach Dixon :

$$\frac{x_{(2)} - x_{(1)}}{x_{(n)} - x_{(1)}} = \frac{18}{22} = 0{,}818$$

Der Prüfwert beträgt 0,818 und ist damit größer als der Schwellenwert von 0,642. $x_{(1)}$ darf also nach dem

Dixon-Test als Ausreißer angesehen und eleminiert werden. Beim korrigierten Ergebnis fallen sofort die deutlich kleineren (d.h. besseren) Werte für s und $T_{\bar{x}}$ auf. Das Ergebnis ist wirklich verbessert worden; mit dem neuen Stichprobenmittelwert wird μ präziser geschätzt.

Beispiel II :

$x_{(1)} = 84$; $x_{(2)} = 90$; $x_{(3)} = 106$; $x_{(4)} = 110$; $x_{(5)} = 110$

$\bar{x} = 100$

$s = 12,17$

$T_{\bar{x}} = 15,12$ (bei beidseitiger Abgrenzung und P=95%)

Auch hier wäre $x_{(1)}$ nach der 15%-Regel zu eleminieren. Das neue Resultat würde lauten :

$\bar{x} = 104$

$s = 9,52$

$T_{\bar{x}} = 15,14$ (bei beidseitiger Abgrenzung und P=95%)

nach Dixon :

$$\frac{x_{(2)} - x_{(1)}}{x_{(n)} - x_{(1)}} = \frac{6}{26} = 0,231$$

Da der Prüfwert 0,231 kleiner als der Schwellenwert 0,642 ist, darf $x_{(1)}$ nicht als Ausreißer behandelt werden. Das Eleminieren dieses Wertes wäre demnach ein Fehler. Bei dem neuen falschen Ergebnis fällt auf, daß s nicht wesentlich kleiner geworden ist. Da n jetzt auch noch abgesunken ist, hat sich das Ergeb-

nis nicht verbessert, sondern im Grunde genommen verschlechtert. Dies objektiviert $T_{\bar{x}}$. Während der Wert beim ersten Ergebnis 15,12 betrug, hat er sich beim zweiten, "korrigierten" Ergebnis aus 4 Einzelwerten auf 15,14 erhöht, also verschlechtert.

Es ist zu erkennen, daß eine Prozentgrenze keine objektive Beurteilung über einen ausreißerverdächtigen Einzelwert zuläßt. Wohl kann in einzelnen Fällen (wie inBeispiel I) der fragliche Einzelwert tatsächlich ein Ausreißer sein. Aber wie in Beispiel II gesehen, geht die Eleminierung eines normalen zufälligen Fehlers mit einer Ergebnisverschlechterung einher, da eine wichtige Information verlorengeht.

2.3.5. Ringversuche

Mit Hilfe einer genügend großen Anzahl von Einzelmessungen in einem Labor gleichen sich die zufälligen Fehler durch die Mittelwertbildung gegeneinander aus. Dieser nun relativ sichere Mittelwert $\bar{x}$ muß aber noch nicht unbedingt dem wahren Wert entsprechen. Der mehr oder weniger laborspezifische, systematische Fehler Δx ist durch eine Prüfung in einem Labor nicht festzustellen. Hierzu sind sogenannte Ringversuche notwendig. Bevor ich deren Prinzip erläutere, ist es sinnvoll, zwei mögliche Versuchsbedingungen als mögliche Grenzfälle voneinander abzugrenzen (nach DIN 1319 Blatt 3, Abschn. 4.5.).

1. Wiederholbedingungen
   Ein Beobachter bestimmt den Meßwert mit ein und demselben Meßgerät nacheinander unter gleichen Arbeitsbedingungen. Unter Wiederholbedingungen sind systematische Fehler nicht erkennbar.

2. Vergleichbedingungen
   Verschiedene Beobachter führen Messungen in verschiedenen Laboratorien unter Verwendung von verschiedenen Meßgeräten der gleichen Bauart durch. In diesem Fall ist die Standardabweichung im allgemeinen größer als unter Wiederholbedingungen, weil unter Vergleichbedingungen die Meßergebnisse auch noch infolge systematischer Fehler "zwischen den Laboratorien" voneinander abweichen.

Zu einem Ringversuch werden nun mehrere Laboratorien zur Untersuchung eines Materials unter Vergleichsbedingungen beauftragt. Man wird von jedem Labor einen Mittelwert mit den Standardkenndaten erhalten. Was kann man nun aus diesen Angaben herauslesen? Dazu 2 einfache Beispiele :

Vorausgesetzt, alle Laboratorien haben mit einer Standardabweichung von gleicher Größenordnung gearbeitet (d.h. ähnlicher Genauigkeit), aber der Mittelwert eines Labors weicht deutlich von denen der anderen Laboratorien ab, so läßt sich schlußfolgern, daß sich in diesem Labor bei jeder Messung ein Fehler von konstanter Größe, also ein systematischer Fehler, eingeschlichen hat. Dies kann ein schlecht geeichtes Meßinstrument, eine falsche Raumtemperatur, eine abweichende Materialverarbeitungsweise etc. etc. sein.

Oder, falls in einem Labor mit weniger Sorgfalt geprüft wurde, so wird sich das in einer relativ höheren Standardabweichung bemerkbar machen. Das Prüfverfahren selbst gestattete ja offensichtlich kleinere s.

Da in der Statistik nicht wie in diesen Beispielen subjektives Abschätzen zu Aussagen führen darf, gibt es auch für Ringversuche objektive Tests mit festgelegten Untersuchungsschritten. In diesem Fall ist es die sogenannte Varianzanalyse, die die verschiedenen Ergebnisse verarbeitet und deren Bewertung erlaubt. Diese in allen ihren Schritten zu erläutern, würde den Rahmen dieser Arbeit sprengen. Dazu wird auf entsprechende Lehrbücher verwiesen (2).

Nur soviel sei hier angedeutet, daß, um zu einem halbwegs aussagekräftigen Ergebnis zu kommen, mindestens 24 Messungen insgesamt notwendig sind (z.B. von 4 Laboratorien je n = 6 Messungen). Man sollte sich bewußt sein, daß dies in der Praxis einen ziemlich erheblichen technischen und zeitlichen Aufwand darstellt. Ob Ringversuche wirklich bei jeder Prüfserie erforderlich sind, bleibt daher die Frage. Eine zu empfehlende Anwendungsmöglichkeit bestünde jedoch, wenn Normen aufgestellt werden, um die einzelnen Prüfvorschriften auf Genauigkeit und Reproduzierbarkeit zu untersuchen. Dadurch könnten Prüfverfahren, die von sich aus hohe Ungenauigkeiten (große Standardabweichungen s) provozieren, schon vor Inkrafttreten der Norm durch präzisere Methoden ersetzt werden.

3. Allgemeine Probleme der zahnärztlichen Materialprüfung

Neben der Auswertung der Einzelwerte ergeben sich im Bereich der zahnärztlichen Materialprüfung bei eingehender Betrachtung noch weitere Probleme

Ausführlich wurde schon auf die Begrenzung der Anzahl der Einzelwerte n eingegangen. Hier gerät man zwangsläufig in einen Gewissenskonflikt. Es muß zwischen dem materiellen und zeitlichen Aufwand und der Notwendigkeit, ein eindeutiges, sicheres und aussagekräftiges Prüfergebnis zu erzielen, abgewogen werden.

Ein besonderes Problem im Fachgebiet der Zahnmedizin entsteht dadurch, daß im allgemeinen von der Industrie Teilprodukte geliefert werden, die sowohl in der zahnärztlichen Praxis wie auch in den Norminstituten gemischt und zum eigentlichen Produkt verarbeitet werden müssen. So wird z.B. kein Phosphatzement geliefert, sondern ein Pulver und eine Flüßigkeit, die miteinander vermischt zu einem Phosphatzement werden. Dies geschieht in der zahnärztlichen Praxis von speziellen Fachkräften, den Helferinnen, im Norminstitut hingegen von Prüfern, die dafür eigentlich nicht ausgebildet sind. Gerade durch die Verarbeitung kann es aber zu gravierenden Eigenschaftsänderungen kommen. Von einer optimalen Verarbeitung ausgehend, können diese Veränderungen nur eine Qualitätsminderung bewirken. Demnach werden bei der Prüfung nicht

nur die Ausgangsprodukte bewertet, sondern zusätzlich noch die Qualität der Verarbeitung.

Als weiteres Beispiel für die komplexe Problematik sei darauf hingewiesen, daß die Probekörper im allgemeinen aus einer oder 2 Handelspackungen hergestellt werden. Die Ergebnisse aus diesen 1 oder 2 Packungen werden auf eine ganze Produktionsserie übertragen. Bei einer derart kleinen Probennahme sind Rückschlüsse jedoch im Grunde unzulässig.

Die Problematik der "Übertragbarkeit in die Klinik" soll als letztes Beispiel näher erläutert werden. Man bemüht sich zwar in den Normen durch Vorschriften über Raumtemperaturen oder Lagerungen der Probekörper in Öfen, in denen ein "Mundklima" herrscht, die Klinik nachzuahmen, dies kann aber nie 100%-ig gelingen. Z.B. werden für die Druckfestigkeitsprüfung bis zu 24mm hohe zylindrische Prüfkörper zwischen zwei Metallplatten gebracht und auf Druck belastet. Im Zahn werden die Füllungsmaterialien jedoch bei einer einfachen okklusalen Kavität durch die Seitenwände zusätzlich gestützt. Bei mehrflächigen Füllungen entstehen andererseits vermehrte Belastungen durch Stufen und Kanten. Kann diesen komplexen klinischen Anforderungen ein einfacher zylindrischer Probekörper gerecht werden? Ist überhaupt das härteste Material gleichzeitig das beste? Haben Unterschiede von 10, 20 oder 50% irgend eine klinische Relevanz?

Man muß wohl die Materialprüfung mehr als eine Sammlung von Indizien für eine klinische Brauchbarkeit betrachten. Sie sollte dazu dienen, Materialien, die offensichtlich unbrauchbar sind, vom Markt zu entfernen.

Keinesfalls wird man aber mit der Materialprüfung beispielsweise eine Qualitätsrangliste für die Praxis erstellen können.

4. Normvorschriften zur Prüfung von Dental-Werkstoffen

Wie auf fast allen Gebieten der Industrie sind auch im Bereich der zahnärztlichen Materialkunde Normen aufgestellt worden. Sie beinhalten Mindestanforderungen an die verschiedenen Materialien. Der Verbraucher soll vor ungeeigneten Produkten geschützt werden. In diesen Normen werden u.a. Vorschriften zur Prüfung zahnärztlicher Materialien z.B. auf Verarbeitungszeit, Abbindezeit, Löslichkeit, Opazität, Druckfestigkeit etc. gegeben.

Für viele dieser Eigenschaften genügt nach der Norm eine visuelle Prüfung mit relativ subjektiven Kriterien, wie z.B. für die Konsistenz und Modellierbarkeit von Amalgamen : "Es darf weder krümeln noch schmieren. " (DIN 13904 Abschn. 4.3.2.). Wenn dies auch auf den ersten Blick sehr unwissenschaftlich und subjektiv erscheint, so sind andererseits dies doch genau dieselben Anforderungen, die der Zahnarzt später in der täglichen Praxis an das Amalgam stellt. Aufwendige Prüfverfahren, die mit einem Zahlenwert als Ergebnis abschließen, wären in diesem Fall technisch nur sehr schwierig durchzuführen, und unter dem Resultat könnte sich niemand etwas vorstellen (z.B. Modellierbarkeitskoeffizient xy).
Bei anderen mechanisch-physikalischen Eigenschaften hingegen, wie z.B. dem Creep und Flow der Amalgame, Druck- und Biegefestigkeit u.ä. können entsprechende Messungen durchgeführt werden. Aus den Werten kann sich auch der Zahnarzt später ein objektives Bild

von den Materialeigenschaften machen, zumindest im Vergleich zu Werten anderer Fabrikate. Bei derartigen Prüfungen ist es nun auch möglich, statistische Methoden zur Auswertung anzuwenden.

Entsprechend schreiben die Normen nicht nur die Versuchsanordnung vor, sondern auch, wie die Einzelmessungen zum Endergebnis zusammengefaßt werden müssen. Z.B. wird bei den verschiedenen Untersuchungen des Creep und des Flow der Mittelwert aus 2, selten 3 Einzelmessungen als Endergebnis angegeben. Einen Schritt weiter gehen die Auswertungsvorschriften der Druck- und Biegefestigkeitsprüfungen. Diese werden zunächst vorgestellt. Danach wird aufgelistet, wie oft und in welcher Norm die jeweilige Auswertungsvorschrift angewendet wird.

4.1. Auswertungsvorschriften bei Druck- und Biegefestigkeitsprüfungen in Norm-Blättern

Im folgenden werden internationale, amerikanische, australische, britische, französische und deutsche Normen im Hinblick auf die Auswertungsvorschrift bei Druck- und Biegefestigkeitsprüfungen untersucht. Aus dem Ostblock lagen keine Norm-Blätter über Dental-Produkte vor.

Die Materialien haben die Prüfung bestanden, wenn das Ergebnis über der jeweiligen Mindestanforderung lag.

Auswertungsvorschrift Typ I:
Als Endergebnis wird der Mittelwert von 5 Prüfergebnissen angegeben.

Auswertungsvorschrift Typ II :
Der Mittelwert aus 6 Prüfergebnissen wird als Endergebnis angegeben.

Auswertungsvorschrift Typ III :

Es werden 5 Einzelmessungen durchgeführt. Wenn mindestens 4 der 5 Prüfergebnisse unter dem Normwert liegen, hat das Material die Prüfung nicht bestanden. Wenn mindestens 4 der 5 Prüfergebnisse über dem Normwert liegen, hat das Material die Prüfung bestanden. In den anderen Fällen müssen weitere 10 Probekörper hergestellt und geprüft werden. Das Endergebnis ergibt sich als Mittelwert aller 15 Prüfergebnisse.

Auswertungsvorschrift Typ IV :

Das Endergebnis ist der Mittelwert der Prüfergebnisse aus 3 oder mehr Prüfungen aus einer Serie von 5 Probekörpern. Probekörper, deren Wert mehr als 15% unter dem Mittelwert liegt, scheiden aus, und der Mittelwert wird aus dem verbleibenden Probekörpern errechnet. Scheiden mehr als 2 Probekörper aus der Serie von 5 aus, so ist die Prüfung insgesamt zu wiederholen.

Tabelle 1 :

Überblick über die vorgeschriebenen Auswertungsvorschriften bei Druck- und Biegefestigkeitsprüfungen in verschiedenen Normen
(in Klammern das Jahr der Veröffentlichung der entsprechenden Norm)

	nach Typ :			
	I	II	III	IV
Amalgame				
- Entwurf für DIN 13904 (79)				X
- Entwurf für ISO R 1559 (77)				X
- ADA 1 (70)	X			
- AS T.2. (49)				X
Kunststoffüllungswerkstoffe				
- DIN 13922 (79)				X
- ISO 4049 (78)				X
- ADA 27 (78)				X
- BS 5199 (75)				X
- AS 1278 (73)				X
Zinkphosphatzemente				
- DIN 13903 (72)				X
- ISO 1566 (78)			X	
- FDI 6 (67)				X
- ANSI/ADA 8 (77)				X
- BS 3364 (61)				X
- AS 1186 (72)				X
Carboxylatzemente				
- Entwurf 6 und 7 für ISO			X	
- Entwurf für ADA 61 (79)				X
- AS 2194 (78)				X

	nach Typ : I	II	III	IV
Silikat- und Siliko-phosphatzemente				
- DIN 13916 (74)				X
- DIN 13918 (77)				X
- ISO 1565 (78)			X	
- ISO 3824 (77)				X
- ISO 3851 (77)				X
- FDI 7 (61)				X
- ADA 9 (75)				X
- BS 3365 Teil 3 (72)				X
- AS 1454 (73)				X
Zink-Oxyd/Eugenol-füllungswerkstoffe				
- Entwurf für DIN 13924 Teil 1(76)	X			
- ISO 3106 (74)	X			
- ADA 30 (77)				X
Prothesenbasiskunststoffe				
- ISO 1567 (78)		X		

Abkürzungen

DIN	Deutsche Industrie Normen
ISO	International Organization for Standardization
FDI	Federation Dentaire Internationale
ADA	American Dental Association
ANSI	American National Standards Institute
BS	British Standards
AS	Australian Standards

4.2. Interpretation der Auswertungsvorschriften

Generell fällt bei allen Normen auf, daß als Endergebnis immer nur der arithmetische Mittelwert angegeben werden muß. Angaben über die Standardabweichung s, über die Anzahl der bei der Berechnung der Kennwerte benutzten Einzelwerte und über Ausreißer werden nicht expliziert verlangt.

In den Vorschriften I und II werden einfach die 5 bzw. 6 Einzelergebnisse zur Mittelwertbildung herangezogen. Ausreißerverdächtige Einzelwerte werden nicht weiter untersucht.

Die III. Auswertungsvorschrift läßt ebenfalls mögliche Ausreißer unberücksichtigt. Der Vorgang erscheint zunächst logisch, daß man, falls 4 von den 5 Meßwerten oder alle über oder unter der Normgrenze liegen, das Material akzeptiert oder ablehnt. Das kann auch statistisch gesehen richtig sein, wenn der Streubereich der Einzelwerte relativ klein ist und der Mittelwert deutlich über oder unter der Normgrenze liegt. Ist die Stichprobenstandardabweichung s jedoch relativ groß oder liegt der Stichprobenmittelwert $\bar{x}$ nur wenig über oder unter der Normgrenze, sind Entscheidungen auf Grund einer solchen Auswertungsvorschrift statistisch nicht vertretbar.

Die am häufigsten anzuwendende Auswertungsvorschrift IV versucht nun, Ausreißer zu eleminieren; die Form über eine 15%-Klausel ist jedoch fragwürdig. Ein entsprechendes Beispiel wurde schon im Kapitel 2.3.4. gebracht. Da diese Normvorschrift andererseits noch am ehesten an statistische Vorgehensweisen erinnert, möchte ich eine theoretisch mögliche Meßreihe vorstellen, die an Hand dieser Vorschrift bearbeitet, zu offensichtlich irrigen Ergebnissen führen würde.

4.3. Negativbeispiel einer möglichen Meßreihe und Diskussion

170 N/mm^2 ist die Normgrenze bei der Druckfestigkeitsprüfung von Silikatzementen (hier Zahlen ohne Dimension)

Meßreihe :

$x_{(1)} = 131$; $x_{(2)} = 131$; $x_{(3)} = 132$; $x_{(4)} = 170$; $x_{(5)} = 208$

$\bar{x} = 154,4$

$s = 34,33$

$T_{\bar{x}} = 32,73$ (bei einseitiger Abgrenzung und P=95%)

Nach der 15%-Klausel müssen $x_{(1)}$ und $x_{(2)}$ eleminiert werden; die Grenze liegt bei 131,24. Aus den verbleibenden Werten $x_{(3)}$, $x_{(4)}$ und $x_{(5)}$ ergibt sich folgendes Ergebnis :

$\bar{x} = 170$

$s = 38$

$T_{\bar{x}} = 64,06$ (bei einseitiger Abgrenzung und P=95%)

Mit dem Wert $\bar{x} = 170$ hätte das Material die Prüfung bestanden. Eine einseitige Abgrenzung des Konfidenzbereiches war zu wählen, da von der Norm lediglich eine Mindestdruckfestigkeit gefordert wird. Eine obere Grenze wird nicht festgelegt.

Was man nicht erfährt : Mit der Wahrscheinlichkeit P=95% ist der uns unbekannte Populationsmittelwert μ über dem Wert $\bar{x} - T_{\bar{x}} = 105,94$ zu erwarten. Diese untere Grenze des Konfidenzbereiches liegt weit vom Stichpro-

benmittelwert $\bar{x}$ entfernt. Das Ergebnis von $\bar{x} = 170$ hat kaum eine Aussagekraft. Mit fast gleicher Wahrscheinlichkeit hätte das Ergebnis auch 169 lauten können. Durch die Anweisung, nur Werte, die mehr als 15% vom Mittelwert nach unten abweichen, zu eleminieren, blieb der extremste Wert $x_{(5)}$ unangetastet. Bei statistischer Betrachtungsweise fällt auch auf, daß der Wert $T_{\bar{x}}$ mehr als doppelt so groß geworden ist wie beim ersten Ergebnis. Bei alleiniger Angabe des arithmetischen Mittelwertes im Ergebnis bleibt diese Tatsache verborgen. Mit Hilfe von 3 stark streuenden Einzelmessungen und der obigen Auswertungsvorschrift kann also ein sehr unsicheres Ergebnis ermittelt werden, welches die Materialien in "gut" oder "schlecht" klassifiziert.

Bei einer statistischen Auswertung der Meßreihe wäre zunächst der Wert $x_{(5)} = 208$ ausreißerverdächtig, da die Differenz zwischen $x_{(5)}$ und $x_{(4)}$ größer ist als zwischen $x_{(2)}$ und $x_{(1)}$. Daher wäre bei einer Untersuchung nach Dixon folgendermaßen vorzugehen :

$$\frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}} = \frac{38}{77} = 0{,}494$$

Da der Prüfwert 0,494 kleiner ist als der Schwellenwert 0,642, darf $x_{(5)}$ nicht als Ausreißer betrachtet werden. Alle 5 Beobachtungen müssen zur Ergebnisermittlung herangezogen werden. Der Stichprobenmittelwert $\bar{x} = 154{,}4$. Der einseitig nach unten begrenzte Konfidenzbereich mit der Wahrscheinlichkeit P = 95% beginnt beim Wert 121,67 ($T_{\bar{x}} = 32{,}73$).

Genauere Aussagen über μ ließen sich nur mit Hilfe weiterer ausreißerfreier Einzelmessungen machen, durch ein Mehr an richtiger Information.

Durch Anwendung der Vorschrift IV ist es jedoch zu einer groben Informatiosverfälschung gekommen, die eigentlich eine Norm nicht zulassen sollte.

4.4. Problematik des positiven Fehlers

Die Auswertungsvorschrift IV gibt an, daß nur Einzelwerte, die um 15% vom Mittelwert nach unten abweichen, eleminiert werden sollen. Alle Meßwerte über dem Mittelwert werden also automatisch zur Berechnung von $\bar{x}$ herangezogen. Anscheinend hatte man beim Formulieren der Norm die Ansicht vertreten, daß es keine zufällig zu großen Abweichungen nach oben geben kann. Wenn es bei einer oder mehreren Messungen möglich war, höhere Werte als den Mittelwert zu erprüfen, dann hätte man diese Werte auch bei den schlechter ausgefallenen Einzelergebnissen erzielen können. Verfolgte man diese Denkweise bis in die letzte Konsequenz, so müßte man den höchsten Einzelwert als Ergebnis angeben. Dies steht aber voll der Auffassung und Erfahrung gegenüber, daß sich alle Einzelwerte entsprechend der Gauß`schen Sreuung um den Populationsmittelwert μ verteilen. Man bedenkt nicht, daß jeder dieser Meßwerte durch zufällige negative und positive Fehler verfälscht wurde. Positive Fehlerkomponenten werden nicht mehr als Fehler, sondern als positive Materialeigenschaften interpretiert. Dies ist statistisch gesehen unzulässig. Auch positive Ausreißer sind als überzufällige Extremwerte anzusehen und als solche zu eleminieren.

Die am häufigsten angewendete Auswertungsvorschift IV stellt sich nun aber so dar : Durch eine fragwürdige 15%-Regel werden kleinere Einzelwerte, die vielleicht gar keine Ausreißer sind, aussortiert. Nach Ausreißern über dem Mittelwert wird erst gar nicht geforscht. Nachdem man nun bis zu 2 "schlechte" Werte eleminiert hat, resultiert aus den 3 verbliebenen höheren Werten ein zwangsläufig höherer Mittelwert.

5. Vorschläge

Vor der Entwicklung eines geeigneten Lösungsvorschlages für die Datenauswertung gilt es zunächst, Überlegungen anzustellen und Kriterien zu formulieren, die von der Vorschrift erfüllt werden sollten.

5.1. Ergebnisangabe in Standardkenndaten ?

Eine Ergebnisangabe in Standardform würde von einem Zahnarzt, der sich wahrscheinlich noch nie mit Statistik beschäftigt hat, sehr viel verlangen. Es ist deshalb kaum sinnvoll, neben dem Stichprobenmittelwert $\bar{x}$ auch die Werte der Stichprobenstandardabweichung s und des Streubereiches der Einzelwerte T anzugeben. Hingegen sollte außer $\bar{x}$ auch $T_{\bar{x}}$ im Endergebnis genannt werden, da $T_{\bar{x}}$ s und n zu einer relevanten Größe verarbeitet. Unter dem Konfidenzintervall, das mit Hilfe von $T_{\bar{x}}$ berechnet wird, kann sich auch ein statistisch nicht vorgebildeter Leser etwas vorstellen. Da von der Norm Mindestanforderungen gestellt werden, ist der Konfidenzbereich nur nach unten zu begrenzen. Zur Berechnung von $T_{\bar{x}}$ sind deshalb die t-Werte aus Tabelle II zu verwenden.

Die statistische Sicherheit P sollte bei der Materialprüfung generell 95% betragen, so daß eine spezielle Erwähnung im Ergebnis nicht mehr notwendig wird.

Angaben über die Einzelwerte und eleminierte Ausreißer sollten im Prüfprotokoll gemacht werden.

5.2. Voraussetzungen an einen praktikablen Lösungsvorschlag

Die weitere Fragestellung hat nun zu lauten : Wieviele Einzelmessungen sind erforderlich ?
Viele Einzelprüfungen bedingen einen großen finanziellen, technischen und zeitlichen Aufwand. Wenige Einzelprüfungen mindern unweigerlich die Qualität der Ergebnisse. Die Anzahl der n sollte weiterhin abhängig von der Genauigkeit des Prüfverfahrens sein und, wenn es um die Erfüllung einer Norm geht, von der Differenz des Stichprobenmittelwertes von der Normgrenze. Im einen Fall wären mehr, im anderen weniger Beobachtungen notwendig. Wenn man den Ansprüchen der Praxis gerecht werden will, muß die Anzahl der Einzelmessungen in sinnvollen Grenzen variabel sein. Um ein Minimum an Aussagekraft zu garantieren, sollte n 4 nicht unterschreiten. Im anderen Extrem fordern die materiellen Kriterien eine Beschränkung von n auf maximal 20 Werte.
Wichtig ist auch, wie die Stichproben ausgewählt werden. Da wir ein Produkt und nicht eine bestimmte Produktionsserie beurteilen wollen, müssen die Stichproben aus verschiedenen rein zufällig ausgewählten Produktionsserien entnommen werden. Andernfalls würden in das Ergebnis konstante Abweichungen aus der spezifischen Produktionsserie mit eingehen und das Ergebnis verfälschen.

Wir müssen uns generell bewußt sein, daß die Schwankungen der Meßergebnisse ja nicht lediglich auf Fehler in der Messung, sondern auch auf Schwankungen in der Produktion zurückzuführen sind. Aus diesen zwei Faktoren setzt sich die Stichprobenvarianz zusamman.

Beikleineren Stichprobenumfängen ist ein Ausreißer-Test zur Bewertung der Einzelwerte besonders wichtig. Ein einziger Ausreißer kann dann das Ergebnis schon extrem verfälschen. Nach DIN 53 804 Teil 1 ist für $n \leq 29$ der Dixon-Test zu verwenden. Dieser Empfehlung sollte auch bei der Materialprüfung von Dental-Werkstoffen entsprochen werden.

Das Ziel der Ergebnisauswertung sollte sein, anzustreben, daß nicht , wie bisher in den Normen, der Stichprobenmittelwert $\bar{x}$, sondern der Populationsmittelwert μ über der Normgrenze liegt. Da μ jedoch unbekannt ist, muß er durch $\bar{x}$ geschätzt und mit Hilfe des Vertrauensbereiches des Stichprobenmittelwertes abgeschätzt werden. Wenn wir zu 95% sicher sein wollen, daß der Populationsmittelwert μ über der Normgrenze liegt, so muß sich die untere Grenze des Vertrauensbereiches zu P=95% über der Normgrenze befinden. Die Forderung lautet daher :

$$\bar{x} - T_{\bar{x}} \geq \text{Normgrenze}$$

Da in diesem Fall der Vertrauensbereich lediglich nach unten begrenzt sein muß, ist Tabelle II zur Ermittlung des Studentfaktors t heranzuziehen.

5.3. Lösungsvorschlag

Aus den im vorangegangenen Kapitel formulierten Voraussetzungen stellt sich der Lösungsvorschlag zur Auswertung der Meßergebnisse wie folgt dar :

1. Zur Anwendung gelangen sollte die folgende Verfahrensweise grundsätzlich bei allen Eigenschaftsprüfungen, die mit einem Zahlenwert als Ergebnis abschließen, z.B. Prüfungen der Druck-, Biege- und Abscherfestigkeit, der Löslichkeit, des Creep, des Flow etc.

2. Die statistische Sicherheit P sollte bei Materialprüfungen grundsätzlich 95% betragen.

3. Die betroffenen Laboratorien sollten in ihrem Bemühen um Genauigkeit mit Hilfe regelmäßig durchgeführter Ringversuche die Qualität ihrer Messungen selbst kontrollieren und nach Fehlerquellen forschen.

4. Die Stichproben sollten aus verschiedenen Packungen verschiedener Produktionsserien entnommen werden und in zufälliger Reihenfolge untersucht werden.

5. Es werden 4 Einzelmessungen durchgeführt. Die Werte werden nach Dixon auf Ausreißer untersucht. Müßten ein oder mehr Werte als wahrscheinliche Ausreißer eleminiert werden, so sind in entsprechendem Umfang weitere Einzelmessungen durchzuführen bis 4 ausreißerfreie Werte vorliegen.

6. Es werden der Stichprobenmittelwert $\bar{x}$, die Stichprobenvarianz s^2 und $T_{\bar{x}}$ berechnet. Liegt der Wert $\bar{x} - T_{\bar{x}}$ über der geforderten Normgrenze, so ist die Prüfung abzubrechen und das Material anzunehmen. Ist die geforderte Normgrenze größer als $\bar{x} - T_{\bar{x}}$, ist eine weitere Einzelmessung durchzuführen.

7. Die gesamte Stichprobe von n = 5 muß nun wieder auf Ausreißer geprüft werden. Eleminierte Werte müssen durch neue Messwerte ersetzt werden, die natürlich auch auf Ausreißer zu prüfen sind.

8. Nun ist wieder wie unter Punkt 6. zu verfahren und zu entscheiden.

9. Liegt die geforderte Normgrenze noch innerhalb des Vertrauensbereiches des Stichprobenmittelwertes $T_{\bar{x}}$, ist wie unter Punkt 7. für n = 6 vorzugehen, danach wie unter Punkt 6.

10. Solange die geforderte Normgrenze innerhalb des Vertrauensbereiches liegt, ist abwechselnd nach Punkt 7. und 6. für die jeweils um einen Meßwert größer werdende Stichprobe vorzugehen.

11. Sobald der Stichprobenumfang 20 Werte überschreiten würde, ist die Prüfung abzubrechen. Das Material hat in diesem Fall die Prüfung nicht bestanden, da nicht mit 95%-iger Wahrscheinlichkeit geklärt werden konnte, ob der Populationsmittelwert μ über der Normgrenze zu erwarten ist.

12. Als Ergebnis der Prüfung soll der aus allen ausreißerfreien Beobachtungen ermittelte Stichprobenmittelwert $\bar{x}$ und $T_{\bar{x}}$ angegeben werden.

5.4. Konsequenzen

Welche Auswirkungen hätte nun diese Forderung? Ist es überhaupt möglich, mit auf 20 Messungen begrenzten Untersuchungen diese Sicherheit von 95% zu erreichen? Oder ist es vielleicht schon möglich, z.B. bei einem sehr exakten Prüfverfahren oder bei einem weit über der Normgrenze liegendem Stichprobenmittelwert $\bar{x}$ mit 4 Messungen auszukommen? In diesem Fall ließe sich ein sicheres Ergebnis mit geringem wirtschaftlichem Aufwand erzielen. Dazu möchte ich im weiteren eine denkbare Prüfsituation besprechen.

Nehmen wir an, die Normgrenze für die Druckfestigkeit eines Materials läge bei 170 N/mm^2, wie dies in den Normen für Silikatzemente gefordert wird. Nach 4 Einzelmessungen ergäbe sich ein knapp über der Normgrenze liegender Stichprobenmittelwert $\bar{x} = 175$. Durch den Dixon - Test sei kein Ausreißer festgestellt worden. Bedingt durch ein relativ präzises Prüfverfahren sei die Stichprobenvarianz mit $s^2 = 25$ ($s = 5$) verhältnismäßig klein. Könnte man mit diesen Daten abschätzen, ob der uns unbekannte Populationsmittelwert u ebenfalls über der Normgrenze liegt? Mit Hilfe des Vertrauensbereiches (bei einseitiger Abgrenzung nach unten) läßt sich μ abschätzen :

$$T_{\bar{x}} = \frac{s \cdot t}{\sqrt{n}} = \frac{5 \cdot 2{,}353}{\sqrt{4}} = 5{,}88$$

In 95% aller Fälle ist μ über dem Wert $\bar{x} - T_{\bar{x}} = 169{,}12$ zu erwarten. Da sich der Stichprobenmittelwert $\bar{x}$ über der Normgrenze befindet, ist zwar der Populationsmittelwert μ mit mehr als 50% Wahrscheinlichkeit über der Normgrenze zu erwarten, eine 95%ige Sicherheit ist jedoch noch nicht gegeben, da die untere Grenze des Vertrauensbereiches die Normgrenze noch nicht überschritten hat.

Aus diesem Grunde müßte ein weiterer ausreißerfreier Meßwert gewonnen werden. Da dieser neue Wert mit dem gleichen Prüfverfahren ermittelt werden würde, wären in der Praxis nur geringe Korrekturen der Stichprobenvarianz von $s^2 = 25$ zu erwarten. In meinem Beispiel möchte ich der Einfachheit halber weiterhin von einer Stichprobenvarianz gleichen Betrages ausgehen, ebenso von dem Stichprobenmittelwert $\bar{x} = 175$. Bei $n = 5$ käme es zu den Werten:

$$T_{\bar{x}} = \frac{s \cdot t}{\sqrt{n}} = \frac{5 \cdot 2{,}132}{\sqrt{5}} = 4{,}76$$

Bei unseren rein theoretisch angenommenen Daten $\bar{x} = 175$ und $s^2 = 25$ ($s = 5$) käme es also bereits nach 5 Einzelmessungen zu einem Ergebnis mit mindestens 95%iger Sicherheit. Denn die untere Grenze des Vertrauensbereiches liegt bei dem Wert 170,24, über der geforderten Normgrenze von 170.

Zur Abschätzung der Konsequenzen des Lösungsvorschlages ist die Tabelle 2 erstellt worden. Ihr liegt folgende Überlegung zu Grunde :

Bei einem bestimmten Stichprobenumfang und einer mit Hilfe jener Einzelwerte berechneten Stichprobenstandardabweichung s muß der Stichprobenmittelwert $\bar{x}$ die Normgrenze mindestens um den Betrag von $T_{\bar{x}}$ übertreffen, damit die Forderung $\bar{x} - T_{\bar{x}} \geq$ Normgrenze erfüllt wird. Z.B muß bei einer Prüfreihe aus 20 ausreißerfreien Beobachtungen und einer daraus ermittelten Stichprobenstandardabweichung s = 5 der Stichprobenmittelwert $\bar{x}$ mindestens um 1,93 ($\hat{=} T_{\bar{x}}$) Einheiten höher liegen als die Normgrenze, sonst muß das Produkt abgelehnt werden.

Tabelle 2 :

Werte für $T_{\bar{x}}$ in Abhängigkeit von n und s bei einseitiger Abgrenzung und P = 95%

s \ n	4	7	10	15	20
5	5,88	3,67	2,90	2,27	1,93
10	11,77	7,34	5,80	4,55	3,87
20	23,53	14,69	11,60	9,09	7,73
40	47,06	29,38	23,19	18,19	15,46

Auf Grund der Bedingung, daß nicht mehr nur der Stichprobenmittelwert, sondern $\bar{x} - T_{\bar{x}}$ die Normgrenze übersteigen soll, kommt es zu einer gewissen Normverschär-

fung, die in ihrem Umfang genau dem Betrag von $T_{\bar{x}}$ entspricht. Damit sich diese Normverschärfung nicht zu stark auswirkt, ist es im Lösungsvorschlag gestattet worden, den Stichprobenumfang bis auf 20 Beobachtungen zu erhöhen. Aus der Tabelle ist ablesbar, daß $T_{\bar{x}}$ mit steigendem n sinkt. Andererseits brauchen nicht alle 20 Messungen durchgeführt zu werden, wenn nach wenigen Beobachtungen schon eine Normerfüllung erreicht ist.

Die zweite Variable, die die Normverschärfung beeinflußt, ist die Stichprobenstandardabweichung s. Wie man in der Tabelle 2 sieht, ist die Normverschärfung um so harmloser, je kleiner s wird. Auch aus diesem Grunde sollten sich die Prüfer stets bemühen, kleine Standardabweichungen zu erreichen, d.h. möglichst viele Fehlerquellen auszuschalten. Eine andere Auswirkung hat s auch auf die notwendige Anzahl der Einzelmessungen. Je kleiner s, desto weniger n werden benötigt. Das spart Zeit und Geld. Zur Selbstkontrolle der Laboratorien sollten deshalb regelmäßig Ringversuche organisiert werden. Für jedes Material andererseits einen eigenen Ringversuch anzusetzen, würde einen sehr hohen materiellen Aufwand bedeuten, der nicht unbedingt gerechtfertigt ist. Wie auf S. 16 schon näher aufgezeigt, könnten Ringversuche zusätzlich bei der Auswahl der Prüfverfahren sehr behilflich sein.

Wenn, wie bei den geltenden Normen, nur $\bar{x}$ als Ergebnis zählt, ist die Entscheidung "Bedingung erfüllt" oder "Bedingung nicht erfüllt" beim Zusammenfallen des Stichprobenmittelwertes $\bar{x}$ mit der Normgrenze nur zu 50% sicher. Nach dem Vorschlag $\bar{x} - T_{\bar{x}} \geq$ Normgrenze ist diese Entscheidung immerhin zu wenigstens 95% wahrscheinlich. Dieses Plus an Sicherheit wird zwar mit einer stillen Anhebung der Normgrenze bezahlt, das sollte jedoch zugunsten der Aussagekraft in Kauf genommen werden, Sollte später bei einer Überarbeitung der Norm nach diesem Vorschlag verfahren werden, so könnte durch eine geringe Absenkung der Normgrenze dieser Tatsache Rechnung getragen werden und damit eine eventuell nicht notwendige Erhöhung der Anforderungen vermieden werden.

6. Schluß

Mittelwerte allein zur Grundlage von Aussagen und Entscheidungen zu machen, bedeutet im Extremfall, Sicherheiten, besser Unsicherheiten, von 50% zu gestatten. Besonders in solchen Fällen ist einfach ein Mehr an richtiger Information, an ausreißerfreien Einzelmessungen, erforderlich, um richtige Entscheidungen treffen zu können. Zusammenfassend für die Untersuchung der Auswertungsvorschriften von geltenden Dental-Normen läßt sich sagen, daß sie dem Ziel der Statistik, der richtigen Informationskonzentration, nicht gerecht werden.

Um zu sicheren Ergebnissen zu gelangen, hätten auch noch weitere Verfahren Anwendung finden können. Durch eine bewußte Beschränkung auf einfache Verfahren blieb der Vorschlag jedoch in einem wirtschaftlich vertretbaren Rahmen. Eine aufgeblähte Auswertungsvorschrift zu fordern, ginge zweifellos an der Realität vorbei.

A N H A N G

Tabelle I : Wertetabelle für t
(bei beidseitiger Abgrenzung)

f	P= 95%	P= 99%	P= 99,9%
1	12,71	63,66	636,62
2	4,30	9,92	31,60
3	3,18	5,84	12,92
4	2,78	4,60	8,61
5	2,57	4,03	6,86
6	2,45	3,71	5,96
7	2,37	3,50	5,41
8	2,31	3,36	5,04
9	2,26	3,25	4,78
10	2,23	3,17	4,59
11	2,20	3,11	4,44
12	2,18	3,06	4,32
13	2,16	3,01	4,22
14	2,15	2,98	4,14
15	2,13	2,95	4,07
16	2,12	2,92	4,02
17	2,11	2,90	3,96
18	2,10	2,88	3,92
19	2,09	2,86	3,88
2o	2,08	2,85	3,85
25	2,060	2,787	3,725
30	2,042	2,750	3,646
35	2,030	2,724	3,592
40	2,021	2,704	3,551
45	2,014	2,689	3,521
50	2,009	2,678	3,496
100	1,984	2,626	3,390
200	1,972	2,601	3,340
300	1,969	2,595	3,328
400	1,967	2,590	3,318
500	1,965	2,586	3,310
600	1,964	2,585	3,307
700	1,963	2,584	3,304
800	1,963	2,583	3,302
∞	1,960	2,576	3,291

Tabelle II : Wertetabelle für t
(bei einseitiger Abgrenzung)

f	P= 95%	P= 99%	P= 99,9%
1	6,314	31,82	318,3
2	2,920	6,965	22,33
3	2,353	4,541	10.21
4	2,132	3,747	7,173
5	2,015	3,365	5,893
6	1,943	3,143	5,208
7	1,895	2,998	4,785
8	1,860	2,896	4,501
9	1,833	2,821	4,297
10	1,812	2,764	4,144
11	1,796	2,718	4,025
12	1,782	2,681	3,930
13	1,771	2,650	3,852
14	1,761	2,624	3,787
15	1,753	2,602	3,733
16	1,746	2,583	3,686
17	1,740	2,567	3,646
18	1,734	2,552	3,610
19	1,729	2,539	3,579
20	1,725	2,528	3,552
21	1,721	2,518	3,527
22	1,717	2,508	3,505
23	1,714	2,500	3,485
24	1,711	2,492	3,467
25	1,708	2,485	3,450
26	1,706	2,479	3,435
27	1,703	2,473	3,421
28	1,701	2,467	3,408
29	1,699	2,462	3,396
30	1,697	2,457	3,385
40	1,684	2,423	3,307
50	1,676	2,403	3,261
60	1,671	2,390	3,232
80	1,664	2,374	3,195
100	1,660	2,364	3,174
200	1,652	2,345	3,131
500	1,648	2,334	3,107
∞	1,645	2,326	3,090

Tabelle III : Formeln und Schwellenwerte für den Dixon-Test

n	Schwellenwert für P=95%	Prüfgröße für $x_{(n)}$	Prüfgröße für $x_{(1)}$
3	0,941		
4	0,765		
5	0,642	$\frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}}$	$\frac{x_{(2)} - x_{(1)}}{x_{(n)} - x_{(1)}}$
6	0,560		
7	0,507		
8	0,554	$\frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(2)}}$	$\frac{x_{(2)} - x_{(1)}}{x_{(n-1)} - x_{(1)}}$
9	0,512		
10	0,477		
11	0,576	$\frac{x_{(n)} - x_{(n-2)}}{x_{(n)} - x_{(2)}}$	$\frac{x_{(3)} - x_{(1)}}{x_{(n-1)} - x_{(1)}}$
12	0,546		
13	0,521		
14	0,546		
15	0,525		
16	0,507		
17	0,490		
18	0,475	$\frac{x_{(n)} - x_{(n-2)}}{x_{(n)} - x_{(3)}}$	$\frac{x_{(3)} - x_{(1)}}{x_{(n-2)} - x_{(1)}}$
19	0,462		
20	0,450		
21	0,440		
22	0,430		
23	0,421		
24	0,413		
25	0,406		

Literaturverzeichnis

(1) R. Kaiser / G. Gottschalk
Elementare Tests zur Beurteilung von Meßdaten
B. I. Hochschultaschenbücher Bd. 774

(2) G. Gottschalk / R.E. Kaiser
Einführung in die Varianzanalyse und Ringversuche
B. I. Hochschultaschenbücher Bd. 775

(3) Georg Hartwig
Einführung in die Fehler- und Ausgleichsrechnung
Carl Hanser Verlag, München

(4) Graf / Henning / Stange
Formeln und Tabellen der mathematischen Statistik
Springer Verlag, Berlin Heidelberg New York

(5) DIN 55302 Blatt 1
(6) DIN 55302 Blatt 2
(7) DIN 1319 Blatt 1
(8) DIN 1319 Blatt 3
(9) DIN 2257 Blatt 2
(10) DIN 13922
(11) DIN 13903
(12) DIN 13916
(13) DIN 13918
(14) Entwurf für DIN 13904
(15) Entwurf für DIN 13924
(16) ISO 1565
(17) ISO 1566
(18) ISO 1567
(19) ISO 3106
(20) ISO 3824
(21) ISO 3851
(22) ISO 4049
(23) Entwurf für ISO R 1559

(24) Entwurf 6 und 7 für ISO
(25) FDI 6
(26) FDI 7
(27) ADA 1
(28) ANSI/ADA 8
(29) ADA 9
(30) ADA 27
(31) ADA 30
(32) Entwurf für ADA 61
(33) AS T.2.
(34) AS 1186
(35) AS 1278
(36) AS 1454
(37) AS 2194
(38) BS 3364
(39) BS 3365 Teil 3
(40) BS 5199
(41) DIN 53 804 Teil 1

Lebenslauf

Am 15. Dezember 1954 wurde ich als Sohn des Zahnarztes Dr. Konrad Hessel und seiner Ehefrau Gertraude, geb. Laskosky, in Amberg/Oberpfalz geboren.

Meine Kindheit verbrachte ich in Neuss am Rhein. Dort wurde ich am 13. 4. 1961 eingeschult. Nach 4 Volksschuljahren wechselte ich auf das Theodor-Schwann-Gymnasium, wo ich am 30. 5. 1973 die Reifeprüfung ablegte.

Vom 1.7. 1973 bis zum 30. 6. 1975 leistete ich den Wehrdienst als Soldat auf Zeit. Entlassen wurde ich als Leutnant der Reserve der Fernmeldetruppe.

Zum Wintersemester 1975/76 begann ich das Studium der Zahnmedizin an der Universität Düsseldorf, wo ich nach Vorphysikum (Juli 1976) und Physikum (März/April 1978) im Frühjahr 1981 die zahnärztliche Prüfung ablegte.

Herrn Prof. Dr. H. Rehberg bin ich für die Überlassung des Themas, für Ratschläge und Anregungen zu großem Dank verpflichtet.
Mein Dank gilt ebenfalls Prof. Dr. med., Dipl. Math. R. Repges für seine Unterstützung auf dem Gebiet der Statistik.

Druck: Koges GmbH, Universitätsstraße 77a, 5000 Köln 41, Tel.: 405105